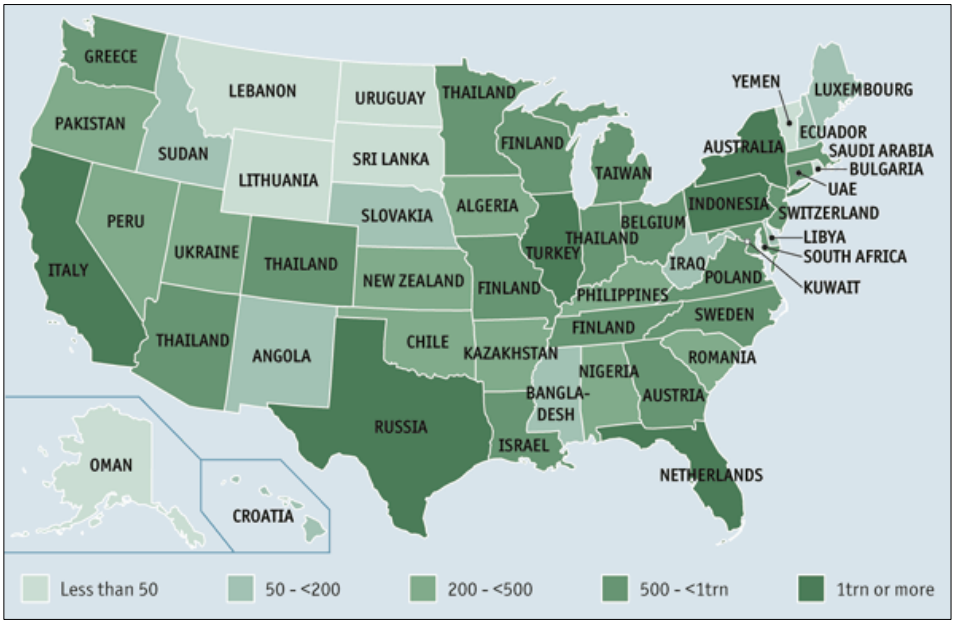
Chapter 1

Introduction

The redrawing of the map comparing U.S. states with countries, shown in Figure 1.1, provides an unusual, yet poignant illustration of how it would be to analyze a cross-section of countries with no explicit reference to where they are located in space or to any of the characteristics of the space in which they are located.

Figure 1.1 Nearest GDP equivalents, 2009 or latest, $bn



Source: Stateside substitutes, *The Economist*, January 13th, 2011

Beyond viewing location as an additional source of information on observations, another essential feature often present in spatially-referenced data is dependence among the observations. At the essence of regional science, economic geography, and

related disciplines is that location and distance matter, resulting in interdependencies in space that influence various forms of economic activity. For example, households may change their location, consumption, and labor supply decisions, and firms may change their location, input demand and output supply decisions, depending on the own economy's market conditions compared to other economies and on the distance to these economies (Elhorst, 2001). The assumption made in standard econometrics that observed values at one location are independent of observations at other locations is thus particularly restrictive and if not addressed, may lead to biased results (see e.g., LeSage and Pace, 2009; Baltagi, 2013).

Irwin and Geoghegan (2001) point out that analyzing spatial data while ignoring interactions among economic agents is analogous to not knowing the chronological order of the observations when analyzing time series data. In time series models, observations closer together in time will generally be more closely related than those further apart. Similarly, the concept of spatial dependence is expressed in Tobler's (1970, p. 236) first law of geography that "everything is related to everything else, but near things are more related than distant things." Despite similarities between temporal dependence and spatial dependence, key distinctions are that the latter is multidirectional and measures can take various forms. These complexities, especially the methodological concerns surrounding the explicit incorporation of interaction effects in regression models is a central focus of the subfield of econometrics known as spatial econometrics.

Since Anselin's (1988) influential book, there has been a flourishing of theoretical and empirical research on the specification and estimation of spatial econometric models, and applications of these methods to study a wide variety of economic and social phenomena. Although too many to mention them all, some contributions include Kelejian and Prucha (1998, 1999), Anselin (2006, 2010), LeSage and Pace (2009, 2011), Elhorst (2010, 2014b), Lee and Yu (2010a, 2010b), Corrado and Fingleton (2012), Arbia and Prucha (2013), and Baltagi et al. (2014), reflecting advances in spatial methods based on cross-sectional data and, since the turn of the century, spatial panels.

This thesis applies and contributes to advances in spatial econometric methods to explore regional labor market dynamics. The main motivations for focusing on regional labor markets in the European Union are threefold. First, labor market

conditions such as employment growth, participation, and unemployment rates are important indicators of economic and social welfare, making the performance of labor markets of great interest to researchers, policymakers, and society at large. Second, substantial and persistent labor market disparities exist across the EU regions. A key point is that labor market variables at the national level can hide enormous differences between regions. Notable examples are provided in Chapters 2 and 4, among which it is shown that unemployment rates within countries can range from around 3 percent in some regions to over 20 percent in others. Third, labor market indicators tend to be correlated across space. This key point is an important example of the observation that economic activity is not randomly distributed across space, which is a shared interest capturing the attention of economists in many fields (see e.g., Brakman et al., 2009). Thus, in addition to taking a regional perspective, we address cross-sectional dependence in the modeling of regional labor markets.

This incorporation of a spatial dimension can make a crucial difference in the analysis and policy implications since the performance of the labor market does not solely depend on factors internal to the regional economy. As emphasized by the OECD (2009), the performance of neighboring regions influences the performance of any other region; regions are closely interconnected through various channels including trade, migration, commuting, capital mobility, and technology diffusion (cf. European Commission, 2014). It should be stressed that since labor market indicators also tend to be strongly correlated over time, we also address serial dependence between the observations on each region over time. We do so by using a dynamic spatial panel data approach, which is exhibiting growing interest in the literature merging time-series, spatial, and panel data econometrics.

One of the studies, with details coming shortly, is a follow-up to a theme issue on spatial econometrics that recently appeared in the *Journal of Regional Science*. Although not on regional labor markets, it is very relevant to the overarching concern of cross-sectional dependence underlying all chapters comprising this thesis. Since researchers often face situations where interdependencies are present (e.g., where economic agents are able to interact), this topic has and continues to be a source of intensive methodological debate. In what follows, each chapter is discussed.

Due to the predicted widening of regional labor market differentials since the recent economic crisis, Chapter 2 focuses on the impact of shocks on regional labor

markets. This issue has gained extensive attention in the literature following the seminal study of Blanchard and Katz (1992) on demand shocks to regional labor markets in the U.S. We explicitly incorporate a spatial dimension by extending their proposed vector autoregressive model of regional unemployment, participation, and employment growth rates to include interaction effects. By allowing for cross-regional interactions, shocks can be transmitted from one region to another, which is more realistic. From the empirical analysis covering 112 regions across the EU over a period of 25 years, we find that besides the direct impact on the region itself, the shock indeed also propagates to other regions, with the impact weakening after around five to seven years on all variables in both the region itself and neighboring regions. A notable difference is that compared to the direct impact, the impact in other regions becomes larger after the first year, reflecting that due to the propagation across space it takes more time for the full effects of the shock to be felt.

It should be emphasized that before carrying out any estimations, crucial methodological concerns are addressed, among which are estimation methods, stationarity, and the specification of the spatial weights matrix W describing the structure of dependence between units in the sample. This is one of the most criticized aspects of spatial econometric models and in this study, several alternative matrices encompassing a wide range of distance and economic distance measures based on empirical and theoretical grounds are considered. Importantly, the search for the best specification of W is done within a system of equations and a Bayesian model comparison approach is used as a selection criterion, in addition to carefully evaluating the stationarity of not only each equation, but also the entire system of equations.

In addition, unlike the original Blanchard and Katz (1992) model, another key contribution of the proposed spatially extended model is that it allows an empirical assessment of the magnitude and significance of so-called spillover effects. Until recently, many empirical studies were interpreting the parameter estimates of spatial models as if they were marginal effects to draw conclusions on whether spillovers exist. However, as pointed out in LeSage and Pace (2009), a change in an explanatory variable of a region will affect the region itself (direct effect) and other regions indirectly (spillover effect) due to complex mathematical formulas. Applying these formulas, it is found that the majority of spillover effects are highly significant,

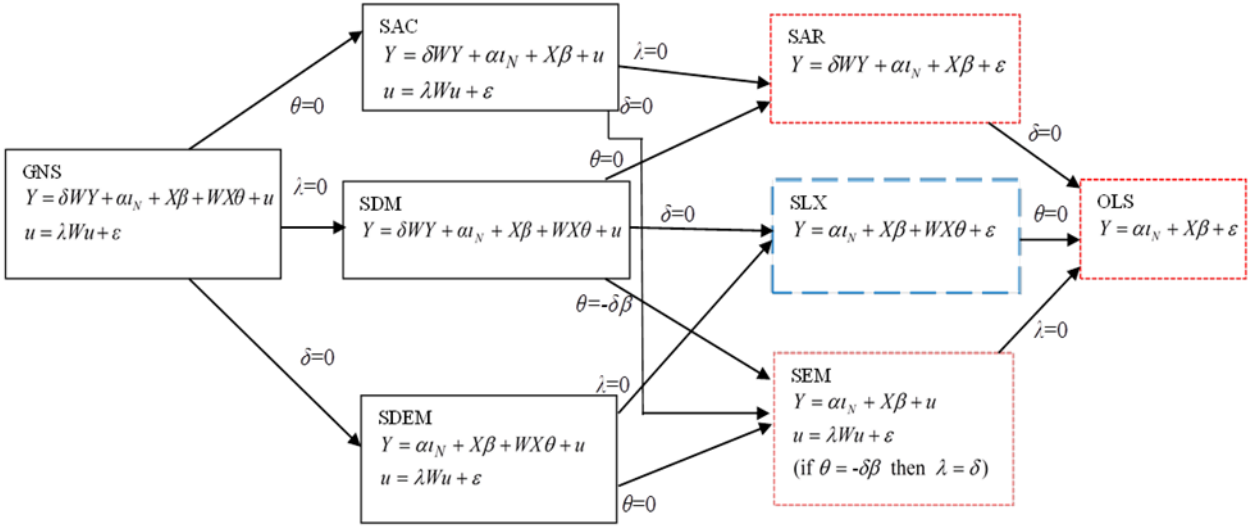
indicating that assuming regions are independent of each other leads to incorrect inference and importantly, a loss of information.

Chapter 3 builds upon the important issue of the interpretation and properties of point estimates compared to marginal effects (the latter consisting of direct and spillover effects) in spatial models discussed in Chapter 2. The origin of this contribution was largely motivated by a recent special issue appraising spatial econometrics in the *Journal of Regional Science* (Partridge et al., 2012). The critical points raised are thought-provoking and led to stimulating discussions, which are reflected in the chapter. As in time-series analysis, this chapter is positioned in chronological order even though it is not on regional labor markets. The main reason for this is that questions on the complex dependencies between units and measuring spillovers that came up in the process, had an impact on the subsequent chapters in the thesis.

Since the main aim of Chapter 3 is to provide more guidance for researchers interested in measuring spillover effects, it can be a useful starting point for readers unfamiliar with spatial econometrics. We provide a comprehensive overview of the strengths and weaknesses of different spatial econometric model specifications in terms of spillover effects. Figure 1.2 presents a helpful diagram to understand the relationships between these different models. Based on this overview, we advocate taking the so-called spatial lag of X (SLX) model as point of departure if a theoretical framework or substantive economic argument indicating which model is most appropriate is lacking. The SLX model, unlike the other models, also allows for W to be parameterized, introducing more flexibility, and the application of standard econometric techniques to test for endogenous explanatory variables.

This starkly contrasts commonly used spatial econometric specification strategies and complements the critique of spatial econometrics raised in the *JRS*. The emphasis on the SLX model is intended as a call for both theoretical and applied spatial econometric work to pay more attention to this model, as well as the issues of whether X and/or WX are endogenous and that W is generally unknown. Distinguishing between global and local spillovers is also highlighted as an important characteristic, with global spillovers implying that a change in X at any location will be transmitted to all other locations including feedback loops, even if two locations according to the specification of W are unconnected. In this respect, limitations of

Figure 1.2 Comparison of different spatial econometric model specifications



Note: GNS = general nesting spatial model, SAC = spatial autoregressive combined model, SDM = spatial Durbin model, SDEM = spatial Durbin error model, SAR = spatial autoregressive model, SLX = spatial lag of *X* model, SEM = spatial error model, OLS = ordinary least squares model (*source* Halleck Vega and Elhorst, 2015a)

widely used global spillover specifications, such as the spatial autoregressive (SAR) model, are stressed. For example, Elhorst (2010) shows that an implication of the spillover properties for this model are that the ratio between the spillover and direct effects is the same for every explanatory variables, which is unrealistic in many empirical analyses. Another issue related to identification problems is the difficulty in discerning whether the significant coefficient of *WY* is due to a causal effect of *WY* or due to omitted, e.g. *WX* variables, making global spillovers more difficult to justify (e.g, see Corrado and Fingleton, 2012).

As an empirical illustration using the Baltagi and Li (2004) cigarette demand model, we find that identification problems are encountered using the standard approach, but are overcome using the SLX model with a parameterized inverse distance matrix which enables estimating the extent at which interaction dampens as distance between units increase. It is found that there is a strong distance decay effect and a significant price spillover effect providing evidence of the well-known bootlegging effect. After dealing with endogenous regressors with IV estimators, consumption appears to have feedback effects on the own-state price, but not for prices in neighboring states, and results change only slightly after controlling for the endogeneity of price. In general, this application demonstrates that ultimately going beyond a traditional viewpoint can introduce more flexibility in modeling interaction effects among regions, potentially generating significant spillovers.

Picking up with this general idea in mind, the following chapters revisit the dynamics of regional labor markets, focusing on labor force participation and unemployment rates. Although unemployment figures are one of the most closely watched and analyzed labor market indicators, Chapter 4 takes a closer look at labor force participation across regions in the EU due to the prevalence of discouragement effects in the wake of the recent crisis. The first contribution is to draw attention to the added value of explicitly considering the spatial dimension. After an extensive perusal of the regional labor market literature, remarkably, only a few studies explore extending the analysis of labor force participation with spatial effects. To gain a better understanding of the mechanisms behind regional dependencies, we draw attention to this issue based on both theoretical and empirical grounds. The motivation to revisit this issue also follows from the critiques on spatial econometrics covered in the previous chapter.

In particular, we propose a shift in focus on interaction effects among the explanatory variables (i.e. an SLX type of model). This is extremely relevant since unemployment, wage, and employment growth differentials (i.e. neighboring labor market conditions) can impact the participation decision in the own-region. Another main issue addressed is the notion of the geographical nature of intergenerational transmission using a time-space recursive modeling approach. Although this approach is rarely used, it can be useful to study spatial diffusion phenomena (Anselin et al. 2008; Korniotis, 2010). This also allows for an appraisal of Fogli and Veldkamp's (2011) recent finding that the participation decision can vary with past participation behavior in surrounding regions (external habit persistence) due to information transmission between people living in different areas.

In addition to variables reflecting local labor market conditions, we also incorporate socio-economic and labor market policy variables which constitute a central part of EU employment strategies, as well as make a gender distinction. In addition, attention is paid to dealing with endogeneity, which is not commonly dealt with in previous regional labor force participation studies and more generally in spatial econometric applications- a point also highlighted in Chapter 3. A system GMM estimator is used to control for endogenous regressors other than the time and space-time lagged dependent variables. Compared to Fogli and Veldkamp's (2011) finding for the US, weak evidence is found for external habit persistence in the EU. Interestingly, interaction effects among the explanatory variables- which has not been previously explored in this context- turn out to be highly significant. Moreover, neglecting spatial effects and potential endogeneity has serious consequences regarding the economic and policy conclusions drawn. This result highlights the importance of the increased attention in the literature on estimating (dynamic) spatial panel models with endogenous variables.

As dependencies among observations are often present in spatially-referenced data, this essential issue itself has been a common thread connecting the chapters of this thesis. Further steps are taken in Chapter 5, building upon and merging past and current developments in the literature. Focusing on regional unemployment rates, three notable stylized facts addressed are that they tend to be strongly correlated over time, parallel the national unemployment rate, and are correlated across space.

Regarding the latter, an intriguing issue to address is the nature of the observed correlation across space.

In the context of regional unemployment rates, a potential source of this correlation can be a result of interactions (e.g. commuting flows) between regions resulting in spillover effects (see e.g., Patacchini and Zenou, 2007). This form of dependence has been extensively covered in the previous studies of the thesis. The observed correlation across space can also be a result of shared factors such as business cycle effects, where outcomes change together as these factors change. Growing attention is being paid in the literature to the difficulty in distinguishing between common factors such as aggregate shocks and actual interactions between units generating spillover effects to explain the observed interdependencies. Although controlling for time fixed effects can help to partly cover common factors, they may not be sufficient. Since the impact of common factors is assumed to be the same across regions, as demonstrated in Chapter 5, this can be restrictive in many applied settings.

Since both sources are likely to be present in applied settings, recent methods have been proposed in the literature to address both forms of cross-sectional dependence, where spatial interaction effects are viewed as 'weak' cross-sectional dependence, and common factors as 'strong' cross-sectional dependence (Chudik et al., 2011). By linking literature on dynamic effects, regional cyclical sensitivity, and spatial econometrics, we are able to address these key observations and propose a methodology that simultaneously accounts for serial dynamics and so-called weak and strong cross-sectional dependence.

It generalizes previous approaches focusing on serial dynamics in combination with weak or strong cross-sectional dependence only, as well as recent approaches employing two-step procedures to account for both types of cross-sectional dependence. We apply these methods using provincial level data for the Netherlands. The substantial and persistent division between high and low unemployment clusters makes it an interesting case and data availability since the early 1970s enables a comparison between prior periods of downturn and recovery to the recent economic crisis. It is shown that the proposed simultaneous approach is more general and that the empirical results produced by previous approaches may lead to biased inference. The implication of this finding is that it is not only relevant to account for common

factors and spatial effects (and serial effects in a dynamic panel setting), but importantly, to also carefully consider the potential consequences of accounting for these effects simultaneously instead of sequentially.

These developments are illuminating, and in combination with the other chapters, the aim has been to take steps forward and add to the discussion surrounding the modeling of intricate dependencies between geographical units. As mentioned in the beginning, understanding the observed non-randomness of economic activity in space is of great interest to economists in various fields, including regional and urban economics. The presence of interdependencies can enrich this understanding, but at the same time, presents significant challenges in the analysis. In sum, Chapters 2, 4, and 5 offer different insights on how this issue can be approached with a focus on regional labor markets, while Chapter 3 takes on board recent criticisms on spatial econometric methods and offers proposals for improved estimation of spillovers, both methodologically and in a conceptual manner. As is common in research, there is no ‘one size fits all’ solution and answers often raise new questions, but this is what makes research exciting and it is hoped that the studies in this thesis contribute to the ongoing debate and promising avenues for further research.